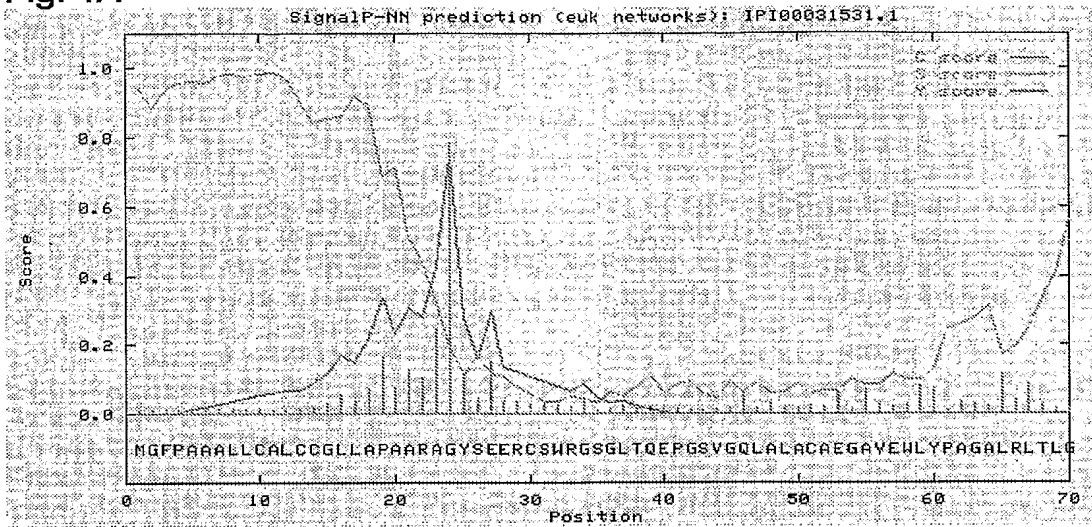


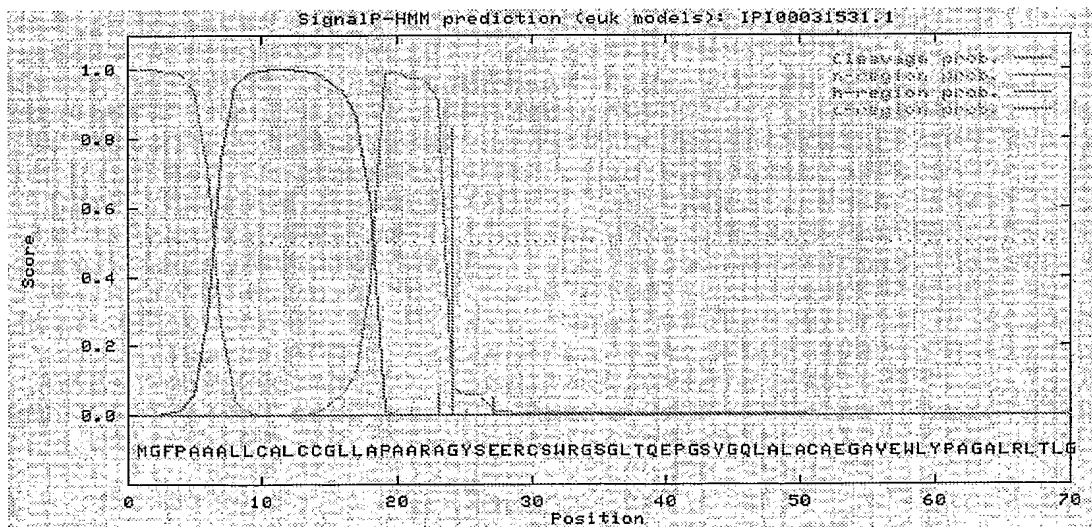
1/13

Fig. 1A



```
>IPI00031531.1          length = 293
# Measure  Position  Value  Cutoff  signal peptide?
  max. C      24      0.785  0.33   YES
  max. Y      24      0.746  0.32   YES
  max. S      11      0.989  0.82   YES
  mean S      1-23    0.846  0.47   YES
# Most likely cleavage site between pos. 23 and 24: ARA-GY
```

Fig. 1B



```
>IPI00031531.1
Prediction: Signal peptide
Signal peptide probability: 1.000
Signal anchor probability: 0.000
Max cleavage site probability: 0.832 between pos. 23 and 24
```

2/13

FIG. 2**human NsG33 (SEQ ID No 3)**

# Gene Ontology category	Odds
Signal_transducer	0.538
Receptor	0.433
Hormone	1.173
Structural_protein	0.168
Transporter	0.230
Ion_channel	0.372
Voltage-gated_ion_channel	0.130
Cation_channel	0.215
Transcription	0.294
Transcription_regulation	0.152
Stress_response	0.340
Immune_response	0.186
Growth_factor	2.083
Metal_ion_transport	0.020

human N-terminal peptide (SEQ ID No 19)

# Gene Ontology category	Odds
Signal_transducer	0.464
Receptor	0.296
Hormone	0.206
Structural_protein	0.987
Transporter	0.311
Ion_channel	0.147
Voltage-gated_ion_channel	0.157
Cation_channel	0.215
Transcription	0.311
Transcription_regulation	0.829
Stress_response	0.162
Immune_response	1.460
Growth_factor	8.142
Metal_ion_transport	0.061

human C-terminal peptide (SEQ ID No 5)

# Gene Ontology category	Odds
Signal_transducer	0.242
Receptor	0.038
Hormone	0.303
Structural_protein	0.096
Transporter	0.231
Ion_channel	0.185
Voltage-gated_ion_channel	0.191
Cation_channel	0.215
Transcription	0.312
Transcription_regulation	0.295
Stress_response	0.145
Immune_response	0.157
Growth_factor	7.963
Metal_ion_transport	0.020

3/13

Fig 3a. CLUSTAL W (1.82) multiple sequence alignment

```

Mouse NsG33      HASAHASALLCALCCGLLAASAHAGYSEDRCSWRGSGLTQEPGSGVGLTLDCTEGAIEWL
Rat NsG33       ---MLVAALLCALCCGLLAASARAGYSEDRCSWRGSGLTQEPGSGVGLTLDCTEGAIEWL
Human NsG33      -MGFPAAALLCALCCGLLAPARAGYSEERCSSWRGSGLTQEPGSGVGLLALACAEGAVEWL
                  .:*****.:*:*****:*****:*****:* *:***:***

Mouse NsG33      YPAGALRLTLGGDPGTRPSIVCLRPFRPFAGAQVFAERMTGNLELLLAEGPDLAGGRCM
Rat NsG33       YPAGALRLTLGGSDPGTRPSIVCLRPTRPFAGAQVFAERMAGNLELLLAEGQGLAGGRCM
Human NsG33      YPAGALRLTLGGDPDPRARPGIACLRPVRPFAGAQVFAERAGGALELLLAEGPGPAGGRCV
                  *****.*:*.*.**** ***** * ***** . *****:

Mouse NsG33      RWGPRRRALFLQATPHRDISRRVAAFREFELHEDQRAEMSPQAQGLGVDGACRPCSDAEL
Rat NsG33       RWGPRRRALFLQATPHRDISRRVAAFQFELHEDQRAEMSPQAQGFVDGACRPCSDAEL
Human NsG33      RWGPRRRALFLQATPHQDISRRVAAFREFELREDGRPELPPQAHGLGVDGACRPCSDAEL
                  *****:*****:***:* *.*:.**:*.*****:*****

Mouse NsG33      LLAactsdfVIHGtiHVAHDTELQESVITVVVARVIRQTLPLFKESGSEGGQGRASIRTL
Rat NsG33       LLTACTSDFVIHGtiHGVVHDMELQESVITVVATRVIROTLPLFQESGSEGRGQASVRTL
Human NsG33      LLAactsdfVIHGtiHGVTHDVELQESVITVVAARVLRQTPPLFQAGRSGDQGLTSIRTP
                  *:***** *****.* *****.:*:*** ***: * * .:* :*:**

Mouse NsG33      LRCGVRPGPGSFLFMGWSRFGEAWLGCAPRFQEFSSRVYSAALTTHLNPCEMALD
Rat NsG33       LRCGVRPGPGSFLFMGWSRFGEAWLGCAPRFQEFSSRVYSAALAAHLNPCEVALD
Human NsG33      LRCGVHPGPGTFLFMGWSRFGEARLGCAPRFQEFRRAYEAAARAHLHPCEVALH
                  *****:***** ***** *.*.*:.**:***:*.

```

Fig 3b. CLUSTAL W (1.82) multiple sequence alignment

```

Mouse      --MLVATLLCALCCGLLAASAHAGYSEDRCSWRGSGLTQEPGSGVGLTLDCTEGAIEWLY 58
Rat        --MLVAALLCALCCGLLAASARAGYSEDRCSWRGSGLTQEPGSGVGLTLDCTEGAIEWLY 58
Human      MGFPAAALLCALCCGLLAPARAGYSEERCSSWRGSGLTQEPGSGVGLLALACAEGAVEWLY 60
            .:*****.:*:*****:*****:*****:* *:***:***

Mouse      PAGALRLTLGGDPGTRPSIVCLRPFRPFAGAQVFAERMTGNLELLLAEGPDLAGGRCMR 118
Rat        PAGALRLTLGGSDPGTRPSIVCLRPTRPFAGAQVFAERMAGNLELLLAEGQGLAGGRCMR 118
Human      PAGALRLTLGGDPDPRARPGIACLRPVRPFAGAQVFAERAGGALELLLAEGPGPAGGRCVR 120
            *****.*:*.*.**** ***** * ***** . *****:

Mouse      WGPERRRALFLQATPHRDISRRVAAFREFELHEDQRAEMSPQAQGLGVDGACRPCSDAELL 178
Rat        WGPERRRALFLQATPHRDISRRVAAFQFELHEDQRAEMSPQAQGFVDGACRPCSDAELL 178
Human      WGPERRRALFLQATPHQDISRRVAAFREFELREDGRPELPPQAHGLGVDGACRPCSDAELL 180
            *****:*****:***:* *.*:.**:*.*****:*****

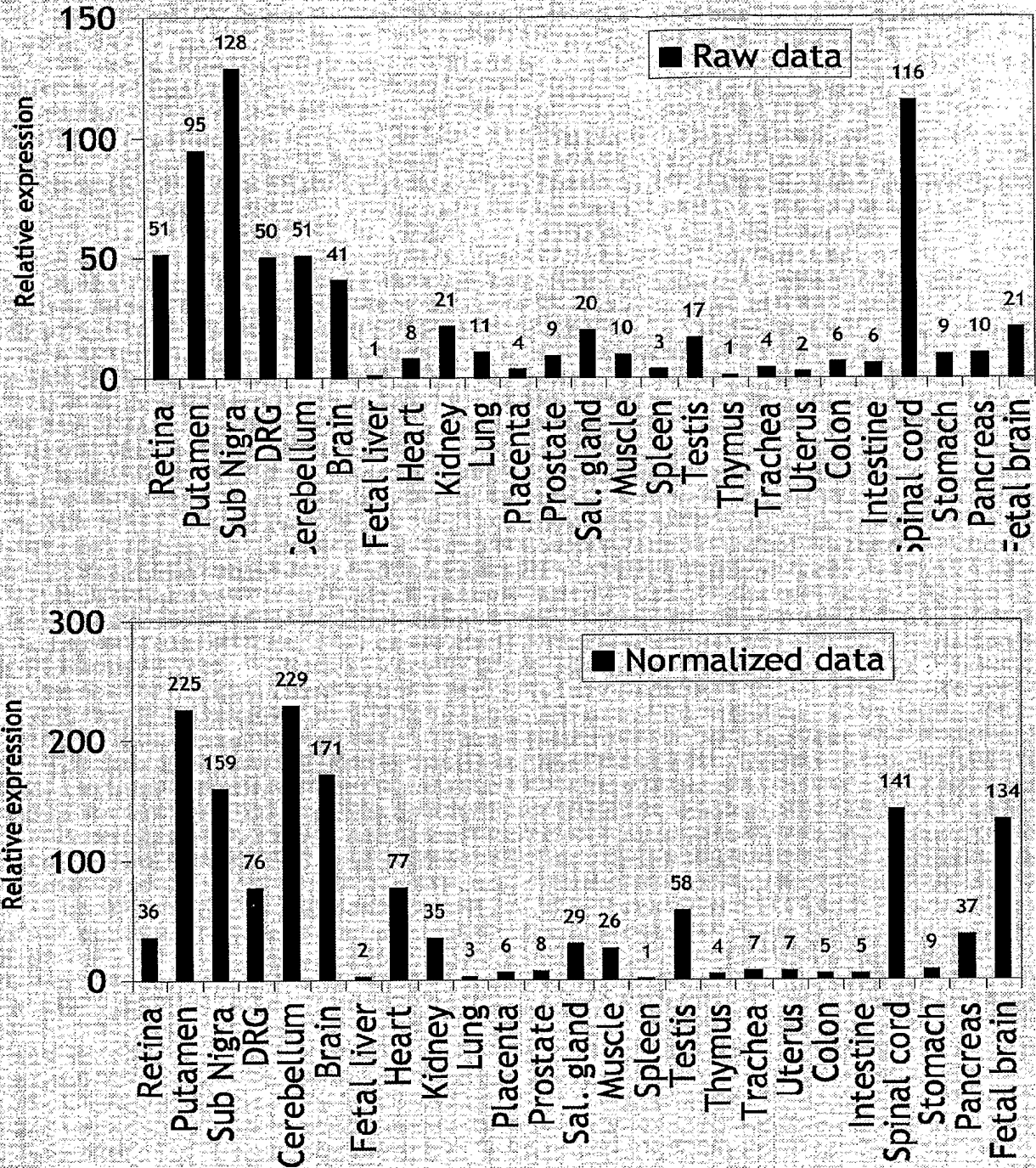
Mouse      LLAactsdfVIHGtiHVAHDTELQESVITVVVARVIRQTLPLFKESGSEGGQGRASIRTL 238
Rat        LTACTSDFVIHGtiHGVVHDMELQESVITVVATRVIROTLPLFQESGSEGRGQASVRTLL 238
Human      LLAactsdfVIHGtiHGVTHDVELQESVITVVAARVLRQTPPLFQAGRSGDQGLTSIRTPL 240
            *:***** *****.* *****.:*:*** ***: * * .:* :*:**

Mouse      RCGVRPGPGSFLFMGWSRFGEAWLGCAPRFQEFSSRVYSAALTTHLNPCEMALD 291
Rat        RCGVRPGPGSFLFMGWSRFGEAWLGCAPRFQEFSSRVYSAALAAHLNPCEVALD 291
Human      RCGVHPGPGTFLFMGWSRFGEARLGCAPRFQEFRRAYEAAARAHLHPCEVALH 293
            *****:***** ***** *.*.*:.**:***:*.

```

4/13

Fig. 4



5/13

scoring matrix: BLOSUM50, gap penalties: -12/-2
 42.3% identity; Global alignment score: 747

```

      10      20      30      40      50 *
Innog. MRGAARAAWGRAGQPWPRPPAPGPPPPPLLLLLLLLAGLLGGAG-AQYSSDRCSWKGSGL
      :      :      :      :      :      :      :      :      :
NsG33  -----MGFPAAALLCALCCGLLAPARAGYSEERCSWRGSGL
              10      20      30

      60      70      *      80      90      100      *110
Innog. THEAHRKEVEQVYLRC AAGAVEWMYPTGALIVNLR-PNTFSPARHLTV CIRSFTDSSGAN
      :      :      :      :      :      :      :      :      :
NsG33  TQEPGS--VGQLALACAEGAVEWLYPAGALRLTLGGPDPR--ARPGIACLRPVRPFAGAQA
      40      50      60      70      80      90

      120     130     140      *      150     160     170
Innog. IYLEKTG-ELRLVPDGDGRPGRVQC--FG-LEQGGLFVEATPQQDIGRRTTG FQYELVR
      .. : : : : : : : : : : : : : : : : : : : : : : : : : :
NsG33  VFAERAGGALELLLAEGPG-PAGGRCVRWGPERRALFLQATPHQDISRRVA AFRFELRE
      100     110     120     130     140     150

      180     190     200     210     220
Innog. RHRAS---DLHEL SAP--CRPCSDTEVLLAVCTSDFAVRGSIQQVTHEPERQDSAIHLRV
      :      :      :      :      :      :      :      :      :
NsG33  DGRPELPPQAHLGLGVDGACRP CSDAELL LAACTSDFVIHGIIHGVT HDVELQESVITVVA
      160     170*      *      180      *      190     200     210

      230     240     250     260     270     280
Innog. SRLYRQKSRVFEPVPEGDGHWQG--RVRTLLECGVRPGHGDFLFTGHMHFGEARLG CAPR
      ... : : : : : : : : : : : : : : : : : : : : : :
NsG33  ARVLRQTPLPLFQAGRSGD---QGLTSIRTPLRCGVHPGPGTFLFMGWSRFGEARLG CAPR
      220     230     240      *      250     260      *

      290     300      * 310
Innog. FKDFQRM YRDAQERGLN PCEVGT D
      : : : : : : : : :
NsG33  FQEFRRAYEAARA AHLHPCEVALH
      270     280     290

```

Fig 5

6/13

```

1  gcttgcgcg ggccgggagg ccggcgcccc cggctgtctc cgccgcgcgc cggaccgcgc ccccgccggg gcagcgggtg tgagagcccc gactccccgg
101  acgcgcgcgc cagtgcacat gggttccccg ccggggcgct gctctgcgcg ctgtgtctgc gactcctggc cccggctgcc cgcgcgggt actccagga
    >>.....CDS.....>
    m g f p a a a l l c a l c c g l l a p a a r a g y s e
201  ggcgtgcagc tggaggggca gggcctcac ccaggagccc ggcagcgtgg ggcagctggc cctggcctgt gcggagggcg cggttgagt gctgtacccg
    >.....CDS.....>
    e r c s w r g s g l t q e p g s v g q l a l a c a e g a v e w l y p
301  gctggggcgc tgcgcctgac cctggggcgc cccgatccca gagcggggcc cggcatcgcc tgtctgcgc cggctgcgc cttgcgggc gcccaggtct
    >.....CDS.....>
    a g a l r l t l g g p d p r a r p g i a c l r p v r p f a g a q v
401  tcgcggagcg cgcagggggc gccctggagc tgctgtggc cgaggggccc ggccggcag ggggccgctg cgtgcgctg ggtccccgc agcgcgggc
    >.....CDS.....>
    f a e r a g g a l e l l l a e g p g p a g g r c v r w g p r e r r
501  cctcttctcg caggccacgc cgcaccagga catcagccgc cgcgtggcgc ccttcgctt tgagctgcgc gaggacgggc gccccgagct gccccgcgag
    >.....CDS.....>
    a l f l q a t p h q d i s r r v a a f r f e l r e d g r p e l p p q
601  gccacggtc tggcgtaga cggtcctgc aggcctgca gcgacgtga gctgtcctg gcgcctgca ccagcgactt cgtaattcac gggatcatcc
    >.....CDS.....>
    a h g l g v d g a c r p c s d a e l l l a a c t s d f v i h g i l
701  atggggtcac ceatgacgtg gagctgcagg agtctgtcat cactgtggtg gccgccgctg tcctccgca gacaccgcg ctgttcagg cggggcgatc
    >.....CDS.....>
    h g v t h d v e l q e s v i t v v a a r v l r q t p l f q a g r
801  cggggaccag gggctgacct ccattcgac cccactgcg tgtggcgtcc acccgggccc aggcaccttc ctcttcattg gctggagcgc cttgggggag
    >.....CDS.....>
    s g d q g l t s i r t p l r c g v h p g p g t f l f m g w s r f g e
901  gcccggtgg gctgtgccc acgattccag gattccgc gtgcctaca ggtgcccgt gctgcccac tcacccctg cgaggtggcg ctgcactgag
    >.....CDS.....>
    a r l g c a p r f q e f r r a y e a a r a a h l h p c e v a l h -
1001  gggctgggtg ctggggaggg gctggtagga gggagggtg gccccactgt ttggaggtga tgggactatc aataagaact ctgttcacgc aaaaaaaaaa
1101  aaaaaaaaaa

```

Fig 6

7/13

```

1  ccacgcgtcc gccacgcgt ccgcgccttct ttgcgcgctc tgttgcggcc tcttgccgcg gtccgctcac gctggctact cggaagaccg ctgcagctgg
>>.....CDS.....
   h a s a h a s a l l c a l c c g l l a a s a h a g y s e d r c s w
101 aggggcagcg gttgaccca ggagcctggc agcgtggggc agctgaccct ggactgtact gagggcgcta tcgagtggct gtaccagct ggggcgctgc
>.....CDS.....
   r g s g l t q e p g s v g q l t l d c t e g a i e w l y p a g a l
201 gctgaccct gggcgccccc gatccgggca cacggcccag catcgtctgt ctggcccag agcgccctt cgctgggtgc caggtcttcg ctgaacgtat
>.....CDS.....
   r l t l g g p d p g t r p s i v c l r p e r p f a g a q v f a e r
301 gacgggcaat ctagagttgc tactggccga gggcccggac ctggctgggg gcgctgcac ggcgtgggt cccgcgcagc gccgagccct ttctctgcag
>.....CDS.....
   m t g n l e l l l a e g p d l a g g r c m r w g p r e r a l f l q
401 gccacaccac accgcgacat cagccgcaga gttgctgcct tccgttttga actgcacgag gaccaacgtg cagaaatgc tccccaggct caaggtcttg
>.....CDS.....
   a t p h r d i s r r v a a f r f e l h e d q r a e m s p q a q g l
501 gtgtggatgg tgctgcagg ccctgcagtg atgccgagct cctcctggct gcctgcacca gtgattttgt gatccacggg accatccatg gggtcgcca
>.....CDS.....
   g v d g a c r p c s d a e l l l a a c t s d f v i h g t i h g v a
601 tgacacagag ctgcaagaat cagtcacac cagtcacac tgtggtggtt gctcgtgtca tcggccagac actgccactg ttcaagggaag ggagctcga gggccaaggg
>.....CDS.....
   h d t e l q e s v i t v v a r v i r q t l p l f k e g s s e g q g
701 cgggcctcca ttctgacctt gctgcgctgt ggtgtgcgtc ctggcccagg ctctctctc ttcatgggtt ggagccgatt tggcgaagct tggctgggtt
>.....CDS.....
   r a s i r t l l r c g v r p g p g s f l f m g w s r f g e a w l g
801 gtgtccccg cttccaagag ttacgcccgtg tctattcagc tgctctcag accatctca accatgtga gatggcactg gactgagaga cctgggagga
>.....CDS.....>>
   c a p r f q e f s r v y s a a l t t h l n p c e m a l d -
901 agccttgat ggaccttctt ctggagatgg ggtgttggg aggtgatgg gaggtgggtt ggaagggtt tggctcggat ggcctcctgg taccacacgt
1001 gagctggtag aatactaagt aatctggacc ataaaaaaaa aaaaaaa

```

Fig 7a

```

1  gggcagccgc gccgcgggct gctcgcgctg cgcgcccgac cctcccgggg cagcagtcgg agggcccgcc gcgtccoccta accatgctgg
   >>CDS.>
   m l
91  tagccacgct tcttgcgct ctctgttgcg gcctcctggc cgggtccgct caogctggct actcggaaga ccgctgcgc tggaggggca
   >.....CDS.....>
   v a t i l c a i c g i l a a s a h a g y s e d r c s w r g
181  ggggtttgac ccaggagcct ggcagcgtgg ggcagctgac cctggactgt actgagggcg ctatcgatg gctgtacca gctggggcgc
   >.....CDS.....>
   s g l t q e p g s v g q i t l i d c t e g a i e w l y p a g a
271  tggcctgac cotggggcgc ccgcatccgg gcacacggcc cagcatgctc tgtctgcgc cagagcgcc ctctcgtggt gccaggtct
   >.....CDS.....>
   i r i t l g g p d p g t r p s i v c l r p e r p f a g a q v
361  tgcgtgaacg tatgaccgc aatctagat tgctactggc cgaggggccg gacctggctg gggggcgcctg catgcgctg ggtccccgcg
   >.....CDS.....>
   f a e r m t g n l e i l i l a e g p d l i a g g r c m r w g p r
451  agggccgagc ccttttctg caggccacac cacaccgga catcagcgc agagttgctg ccttcgctt tgaactgcac gaggaccaac
   >.....CDS.....>
   e r x a l f l q a t p h r d i s r r v a a f r f e l h e d q
541  gtgcagaaat gtctcccgag gctcaagtc ttggtgtgga ttgtgctgc aggcctgca gtgatgcga gctcctctg gctgcacgca
   >.....CDS.....>
   r a e m s p q a q g l g v d g a c r p c s d a e l l i a a c
631  ccagtgttt ttgtatccac gggaccatcc atggggtgc ccatgacaca gagctgcaag aatcagtcac cactgtggtg ttgtcgtg
   >.....CDS.....>
   t s d f v i h g t i h g v a h d t e l q e s v i t v v v a r
721  tcacgcgcca gacactgcca ctgttcaagg aaggagctc ggaggggcaa ggccgggct ccattcgac ctgtcgcgc ttgtgtgtgc
   >.....CDS.....>
   v i r q t l p l f k e g s s e g q g r a s i r t l l r c g v
811  gtccctggcc aggtccttc ctcttcacg gctggagcgg atttgggaa gcttggtgctg gctgtgctcc ccgttccaa gatttcagcc
   >.....CDS.....>
   x p g p g s f l f m g w s r f g e a w l g c a p r f q e f s
901  gtgtctattc agctgtctc acgaccatc tcaaccatg tgagatgga ctggactgag agacctggga gcaagccctg gatggacctt
   >.....CDS.....>
   r v y s a a l t t h i n p c e m a l d -
991  ctctctggaga tgggtgtgtg gggaggggtga tgggaggtg ggtgagaagg gtgtggtctg gatggcatcc tgggtaccac agtgagctg
1081  tagaatacta agtaatctg accataccag ccactgtagt catggtcttc ttgtggcaggc agcatacca gctctgtgc tgcctcactt
1171  tgtctactct ccagtctgct gcccttctaa cccttcttag cctgtgacc agtgagctca tgtttcttc gaattccagg gtgctgctg
1261  ggttcagagc aacgtgccc tagtttgaa gacttgagct aattgtttt tttttgtttg tttttttgtt tgttttaaagg tggcctgggg
1351  gggggggcaa aca

```

Fig 7b

9/13

```

1  atgtgtgtag cggcgcttct ctgcgcgctg tgctggggc tcttggtgc gtccgctcga gctgggtact ccgaggaccg ctgcagctgg aggggcagcg
>>.....CDS.....>
    m l v a a l l c a l c c g l l a a s a r a g y s e d r c s w r g s
101 gttgacca ggaacctggc agctggggc agctgacct ggattgtact gaggtgcta tggagtggct gstatccagct gggcgctgc gcctgactct
>.....CDS.....>
    g l t q e p g s v g q l t l d c t e g a i e w l y p a g a l r l t
201 aggggctct gatcgggca cgcggcccag catcgtctgt ctgcgccaa cagggccctt cgctgggtgc caggtcttcg ctgaacggat ggcgggcaac
>.....CDS.....>
    l g g s d p g t r p s i v c l r p t r p f a g a q v f a e r m a g n
301 ctgagtgc tactggcca gggccaaggc ctggctggg gccgtgcat gcgtgggt cctgcgagc gccgagccct ttctctgcag gccacggcac
>.....CDS.....>
    l e l l l a e g q g l a g g r c m r w g p r e r r a l f l q a t p
401 accgggacat cagccgcaga gttgctgct tocaatttga actgcacgag gaccaacgtg cagaaatgc tccccaggcc caaggttttg gtgtggatgg
>.....CDS.....>
    h r d i s r r v a a f q f e l h e d q r a e m s p q a q g f g v d
501 tgcctgcagg cctgcagtg atgccgagct ccttctgact gcatgcacca gtgactttgt gatccatggg accatccatg gggctgcca tgacatggag
>.....CDS.....>
    g a c r p c s d a e l l l t a c t s d f v i h g t i h g v v h d m e
601 ctgcaagaat cagtcacac tgtgttgcc actcgtgtca tccgccagac actgccactg ttccaggaag ggagctcgga gggccggggc caggcctccg
>.....CDS.....>
    l q e s v i t v v a t r v i r q t l p l f q e g s s e g r g q a s
701 ttctgacctt gttgcgctgt ggtgtgcgtc ctggcccagg ctcttctc tccttctc ttcatgggt ggagccgatt tggcgaagct tggctgggt gcgctcccg
>.....CDS.....>
    v r t l l r c g v r p g p g s f l f m g w s r f g e a w l g c a p
801 ctccaagag ttacgccgtg tctattcagc tgctctcgc gccacactca acccatgtga ggtggcactg gactgagaga cctgggagca agccctggat
>.....CDS.....>>
    r f q e f s r v y s a a l a a h l n p c e v a l d -
901 ggatcttct ctggggatgg ggtgttggg aggggtgata ggagggtggg tgggaagggt gtggctcaga tggcatcctg gtaccacacag tgagggtggtta
1001 gaatactaaa taacctggat cacacc

```

Fig 8

10/13

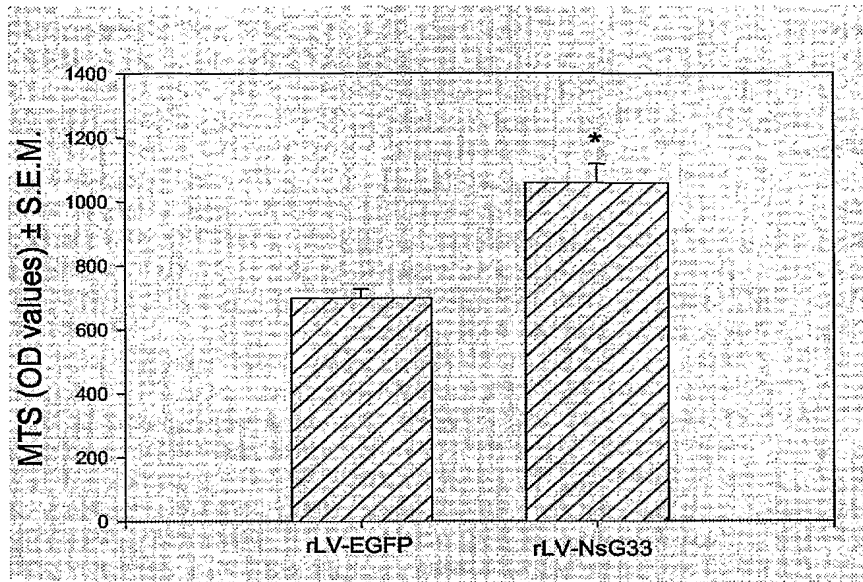


Fig. 9

11/13

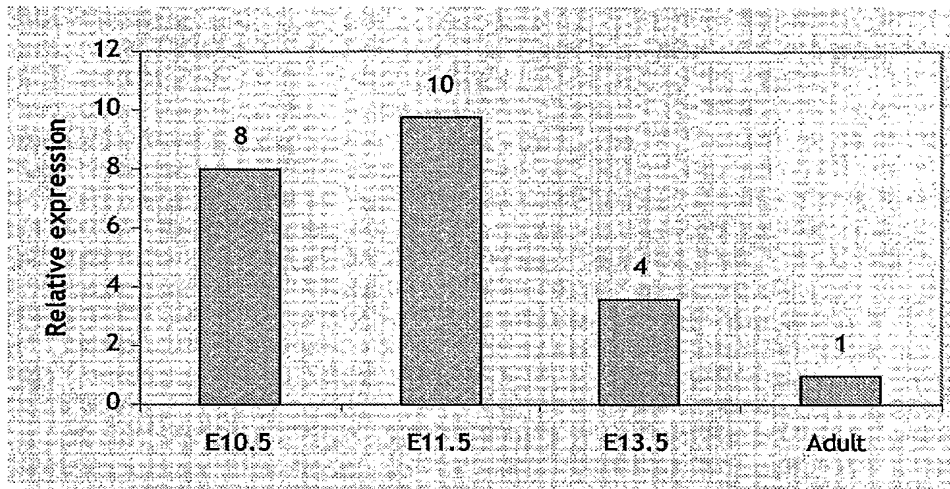


Fig. 10A

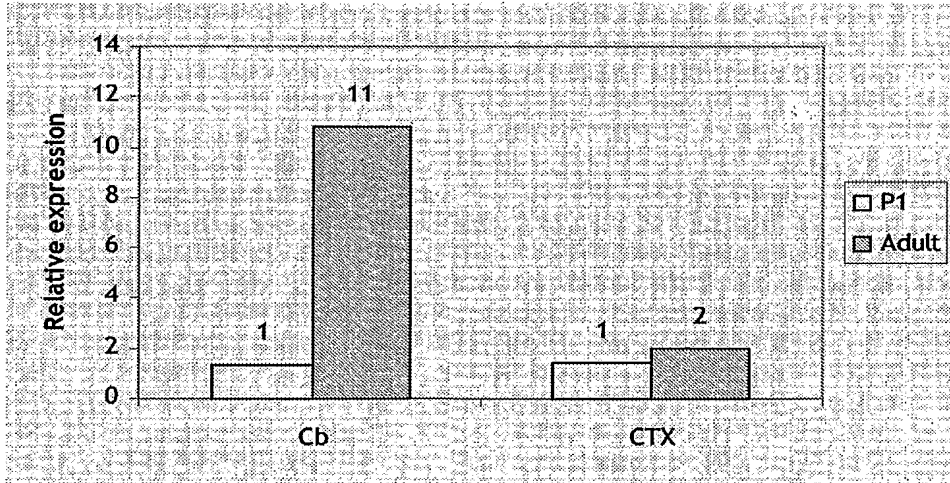


Fig. 10B

12/13

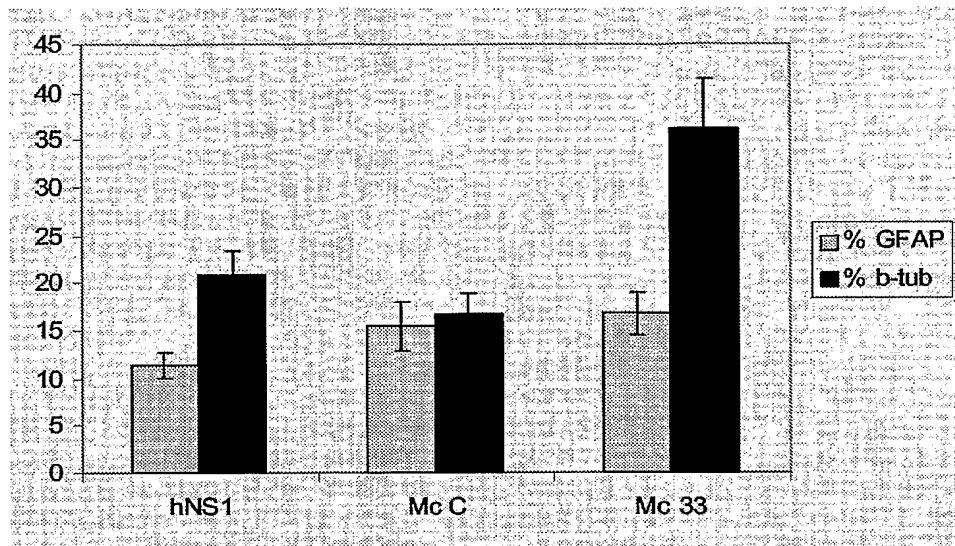


Fig. 11A

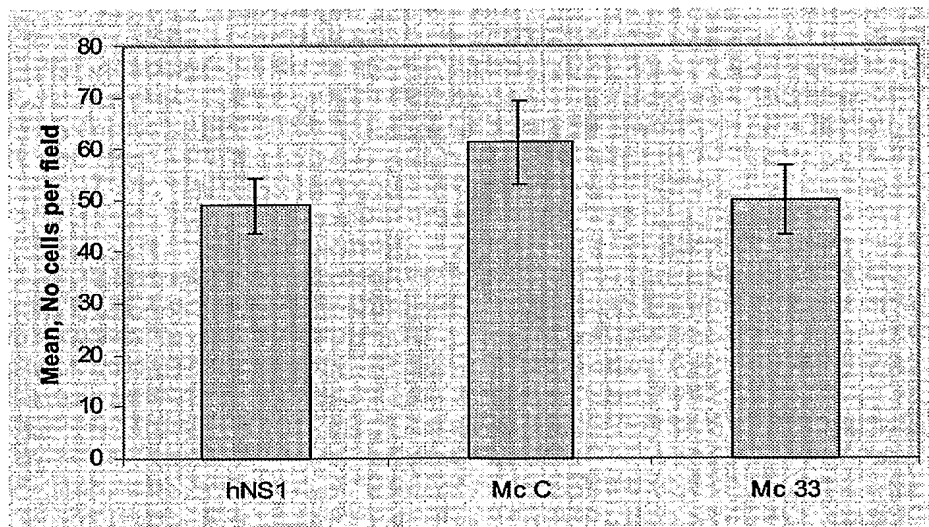


Fig. 11B

13/13

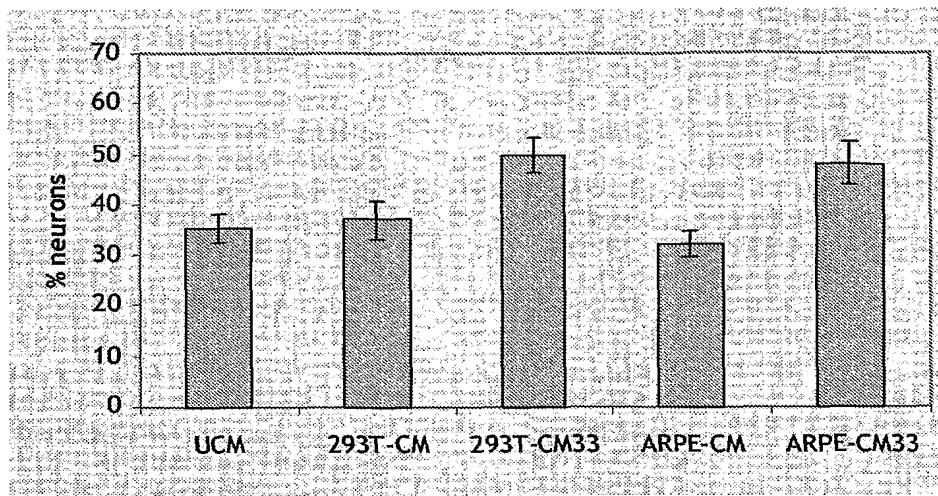


Fig. 12